A Sentiment Analysis of Tourist Attractions in the Mandalika Special Economic Zone (SEZ) Using the Nave Bayes Method

Teguh Iman Pribadi[1*], **Fahry**[2], **Muharis**[3], **Ega Dwi Putri Marswandi**[4]

[1,2,3,4]Universitas Bumigora, Mataram, Indonesia

teguh@universitasbumigora.ac.id[1*], fahry@universitasbumigora.ac.id[2], muharis@universitasbumigora.ac.id[3],
ega_dwp@universitasbumigora.ac.id[4]

Abstract-

Background: The Mandalika Special Economic Zone has become one of the most popular destinations for domestic and international tourists. This popularity highlights the importance of understanding the tourists views and feelings.
Objective: This research aims to classify tourists' opinions regarding the attractions in the Mandalika Special Economic Zone into positive and negative classes.
Methods: This sentiment analysis uses the Nave Bayes classification method. The data analyzed was obtained from 1,144 reviews on the TripAdvisor platform. The research stages included data collection, labeling, preprocessing, transformation, classification, analysis, and visualization.
Result: The analysis yielded 1,144 reviews, with 1,033 reviews having a positive sentiment and 111 reviews having a negative sentiment. The classification results using the Nave Bayes method showed fairly consistent accuracy levels: fold 1: 89.08%, fold 2: 89.96%, fold 3: 88.21%, fold 4: 87.34%, and fold 5: 90.79%.
Conclusion: Based on the data analysis results, most tourists have a positive sentiment towards the attractions in the Mandalika Special Economic Zone. The testing results also show that the Nave Bayes algorithm successfully classified tourist sentiments well, with an average accuracy across all folds of 89.07%.

Keywords: maximum 6 keywords from paper.

Corresponding Author:

Teguh Iman Pribadi,
Study Program of Tourism, Universitas Bumigora, Mataram, Indonesia,
Email: teguh@universitasbumigora.ac.id

1. INTRODUCTION

The rapid development of information technology has encouraged many tourists, especially the millennial generation, to utilize technology to improve their tourism experience [1]. Among the technologies currently developing, social media plays a significant role. Tourists use social media to dig up in-depth information about a destination, which ultimately increases their decision-making awareness and independence. A fundamental part of social media in current technological developments is represented by the dissemination of user-generated content (UGC) [2]. This content includes opinions and comments about purchased products and services or users' life experiences. This content determines the spread of electronic word of mouth (eWOM), which refers to positive or negative comments made by customers regarding products or companies that individuals and institutions can access via the Internet. In tourism, the most important form of eWOM is online reviews written on social networks.

The development of online review platforms has also indirectly changed tourism business trends. Previously, they used traditional business models, but they are now switching to digital business models [3]. With

various features provided, such as reviews, ratings, and photos about various places and services, this online review platform helps users make better decisions in choosing certain places or services and provides opportunities for businesses to improve their reputation and attract more customers.

TripAdvisor is a popular online review platform in the tourism sector. TripAdvisor provides reviews, comments, and ratings about a destination, tourism destinations, accommodations, and restaurants. Recent data shows that TripAdvisor is one of the most visited travel platforms worldwide. TripAdvisor is a company based in Boston, United States, and was founded in 2000. This company is one of the pioneers of electronic word of mouth (eWOM) in the field of travel, which provides various features such as reviews, ratings (1 star 5 stars), photos, and forums related to travel planning, such as tour guides, transportation, tourist attractions, restaurants, and accommodation [4]. Therefore, this website has much information that can be used as a reference for everyone going on a tourist trip.

TripAdvisor has become a very relevant platform in modern tourism due to its success in collecting and presenting reviews from travelers worldwide. The existence of these reviews not only helps other travelers in planning their trips but also provides valuable data that can be analyzed further. One method that can be used to analyze review data is sentiment analysis. Sentiment analysis can be interpreted as a data analysis process by extracting a person's opinions, sentiments, assessments, attitudes, and emotions towards an object expressed in written form [5]. More specifically, sentiment can be divided into two, namely, positive sentiment and negative sentiment. In the tourism sector, this analysis method can provide important information for stakeholders regarding the sentiments, feelings, and emotions of tourists or potential tourists [6].

This research analyzes the sentiments of tourists visiting tourist attractions in the Mandalika Special Economic Zone (SEZ). The Mandalika is located in Central Lombok Regency, West Nusa Tenggara Province. This area is one of five super-priority tourism destinations designated by the Ministry of Tourism and Creative Economy. Many researchers have previously researched the Mandalika, but the research has mostly analyzed the impact of development [7, 8], community empowerment [9–11], and the role of stakeholders [12, 13]. Meanwhile, research on tourist sentiment still needs to be conducted. In fact, sentiment analysis can provide valuable information for policymakers, which can be used as evaluation material for product and service management at a tourism destination [6] .

Several previous studies have analyzed tourist sentiment toward tourist attractions. For example, Utami and Erfina, in 2022, conducted an analysis of tourist sentiment towards tourist attractions in Bali based on reviews on Google Maps [14]. Another research study was conducted by Singgalen in 2022, which analyzed tourist sentiment towards Borobudur Temple based on reviews on TripAdvisor [15]. Apart from being specific to tourist attractions, tourist sentiment analysis can also be carried out on accommodations in a tourist destination. This was done by Ginantra et al. in 2022, who analyzed tourist sentiment towards villas in the Ubud area, Bali [16]. The three studies mentioned above used the same classification method, namely the Nave Bayes classification method. Overall, these studies provide valuable insights for tourism destination and accommodation managers to understand tourist perceptions and identify areas for improvement.

The difference between this research and the research mentioned above, apart from the data source and location, is also a difference in the initial data processing method, where this research uses a word normalizer from prosa.ai, which affects the level of accuracy of the model obtained. With a word normalizer, the retrieved review text data can be processed better, reducing writing errors and language variations commonly found in online reviews. In addition, this word normalizer can change non-standard or misspelled words into a standard form so that the sentiment classification model can work more efficiently and accurately. This research aims to classify tourists' opinions regarding the attractions in the Mandalika Special Economic Zone into positive and negative classes. The results of this research can be used as evaluation material for policymakers to improve destination management, especially in the Mandalika Special Economic Zone.

2. RESEARCH METHODS

This research is quantitative. The resulting data is processed through six stages, starting with data collection and ending with data analysis and visualization. The following flow chart (Figure 1) provides a more detailed explanation of each stage.
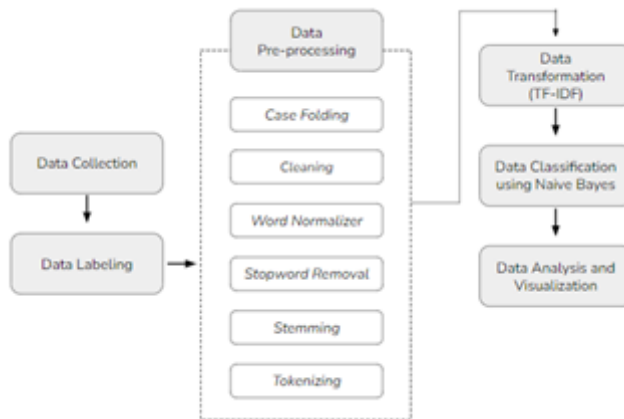

Figure 1. Research Stages

2.1. Data Collection

Data collection is a stage in research that aims to obtain important data or information that will later be analyzed. In this research, the data collected came from the TripAdvisor platform, namely reviews from tourists who visited tourist attractions in the Mandalika. The tourist attractions in question are Kuta Mandalika Beach, Tanjung Aan Beach, Merese Hill, Gerupuk Beach, and Seger Beach. The number of reviews of each tourist attraction can be seen in the following table (Table 1):

Table 1. Kinerja sistem

No	Name of Tourist Attractions	Number of Reviews
1	Kuta Beach	655
2	Tanjung Aan Beach	354
3	Merese Hill	73
4	Gerupuk Beach	11
5	Seger Beach	51
Total Reviews		1.144

2.2. Data Labeling

After getting data from TripAdvisor, labeling was then carried out using the lexicon method. In this method, labeling is carried out on text using words from the lexicon dictionary, which consists of positive and negative words. The score is calculated based on the number of words in each text [17].

If the number of positive words in a text exceeds negative words, the text will be labeled with positive sentiment. Conversely, if the number of negative words in a text exceeds the number of positive words, the text will be labeled with negative sentiment. If a text has the same number of positive and negative words, then the text is labeled neutral.

2.3. Data Pre-processing

The initial data processing in this study was conducted in six stages. The explanation of each stage can be seen as follows. First, Case Folding involves converting every character to lowercase to ensure data consistency [18]. Second, Cleaning aims to remove unwanted characters, hashtags, and URLs to reduce noise in the classification process [**?**]. Third, Word Normalizer corrects misspelled words using the prosa.ai platform.

Fourth, Stopword Removal eliminates irrelevant words such as this, that, and which [19]. Fifth, Stemming extracts the root form of a word using the Phyton Sastrawi library. Finally, Tokenizing separates sentences into individual words to facilitate data transformation.

2.4. Data Transformation

Data transformation means converting existing data in one format into a new format. In this stage, the technique used is Term Frequency-Inverse Document Frequency (TF-IDF). This technique analyzes the presence of words in each document and assigns weight values to them [20]. Term Frequency (TF) is calculated based on the number of occurrences of each word in each document, while Inverse Document Frequency (IDF) is calculated based on the number of occurrences of words in the entire document collection.

2.5. Data Classification

The transformed data is then classified using the Nave Bayes method. This method was discovered by a scientist from England named Thomas Bayes. Nave Bayes is a probability-based classification method that uses Bayes' theorem, assuming that each feature is highly independent. Nave Bayes only requires relatively little training data to estimate the parameters needed in the classification [15].

2.6. Data Analysis & Visualization

After classifying the data using Nave Bayes, the final step in this research is to analyze and visualize the data resulting from the classification. Data analysis aims to interpret classification results and identify significant patterns or trends. Data visualization, on the other hand, aims to present analysis results in graphic or image form that is easy to understand.

3. RESULTS AND DISCUSSIONS
3.1. Data Collection

As previously informed, the data analyzed in this research are online reviews of tourists on the TripAdvisor platform, totaling 1,144 reviews. A web scraping technique was used to make data collection easier. Web scraping is the process of collecting data from websites using a specific programming language, in this case, Python.

Data obtained from web scraping results include the reviewer's name, rating from 1-5, tags or keywords, review content, and review date. Examples of web scraping data can be seen in the following table (Table 2):

Table 2. Example of Data Scraping

review_id	review_name	review_rating	review_tag	review_content	review_date
1	Dayu Kartika	5	Liburan 3 hari ke Lombok	Pantai Kuta Mandalika, pantainya indah, pasir...	2019-06-27
2	Wardhanie	4	Pantai yg eksotis	Pantai berpasir putih dan bersih, ada ayunan, mi...	2019-06-12
3	Duty Nath	5	Pantai yang ramai	Pantai Kuta adalah pantai yang penuh keramaian...	2019-06-01
4	Andy Lombok Transport	4	Tempat bagus di Lombok	Pantai Kuta Lombok tempat yang tepat untuk dud...	2019-05-27
5	UD Talita	5	Indah asri aman nyaman	Pantainya sejuk, udaranya ademmmm Rekomend b...	2023-05-23

3.2. Data Labeling

After collecting all the review data, the next step is translating the review content from previously used Indonesian or other languages to English (Table 3). Translation into English is the initial stage before labeling data using the lexicon method. This is done because the lexicon will work better if it uses English, while other languages, such as Indonesian, are not yet available in the lexicon library.

Table 3. Translation Table

review	translate	lexicon	
0	Pantai kuta mandalika, pantainya indah, pasir...	Kuta Mandalika Beach, the beach is beautiful, ...	positive
1	Pantai berpasir putih dan bersih, ada ayunan, mi...	The beach is white sand and clean, there are swings, mi...	positive
2	Pantai kuta adalah pantai yang penuh keramaian...	Kuta Beach is a beach full of crowds. Around the beach...	positive
3	Pantai kuta lombok tempat yang tepat untuk dud...	Kuta Lombok Beach is the right place to sit back...	positive
4	Pantainya sejuk, udaranya ademmmm Rekomend b...	The beach is cool, the air is cool Highly recommended b...	positive

The results of labeling 1144 reviews using the lexicon method found that there were 1,033 positive labels and 111 negative labels. Positive labels are a classification given to text or data containing sentiments considered positive. For example, reviews that praise, praise, or express satisfaction. On the other hand, a negative label is a classification given to text or data that contains sentiments that are considered harmful. For example, reviews that criticize, express disappointment or provide complaints.

3.3. Data Pre-processing

Pre-processing must be carried out beforehand to facilitate the data processing process. This process is carried out in five stages, namely changing each letter character to lowercase (case folding); cleaning unwanted data (cleaning); correcting words that are miswritten (word normalizer); removing words that are considered irrelevant in a sentence (stopword removal); taking the base word of a word (stemming), and finally separating sentences in a text into words (tokenizing). An example of the pre-processing stage can be seen in the following table (Table 4).

Table 4. Example of Pre-processing Stage

review	lexicon	lowercase	clean	normal	stopword	stemmed	token
Pantai kuta mandalika, pantainya indah, pasir...	positive	pantai kuta mandalika , pantainya indah , pasir...	pantai kuta manda- lika pan- tainya indah pasirnya...	pantai kuta pan- tainya indah pasirnya...	pantai kuta pan- tainya indah pasirnya...	pantai kuta mandalika pantai indah pasir besar...	[pantai, kuta, mandalika, pantai, indah, pasir...]
Pantai berpasir putih dan bersih, ada ayunan, mi...	positive	pantai berpasir putih dan bersih , ada ayunan , mi...	pantai berpasir putih dan bersih ada ayunan mi...	pantai berpasir putih dan bersih ada ayunan mi...	pantai berpasir putih bersih ayun mini tempa...	pantai pasir putih bersih ayun mini tempat...	[pantai, pasir, putih, bersih, ayun, mini, tempat...]
Pantai kuta adalah pantai yang penuh keramaian...	positive	pantai kuta adalah pantai yang penuh keramaian...	pantai kuta adalah pantai yang penuh keramaian...	pantai kuta pantai penuh keramaian sekitar pan...	pantai kuta pantai ra- mai sekitar pantai...	pantai kuta pantai ra- mai sekitar pantai...	[pantai, kuta, pantai, penuh, ra- mai, sekitar, pantai...]

review	lexicon	lowercase	clean	normal	stopword	stemmed	token
Pantai kuta lombok tempat yang tepat untuk dud...	positive	pantai kuta lombok tempat yang tepat untuk dud...	pantai kuta lombok tempat yang tepat untuk dud...	pantai kuta lombok tempat tepat duduk santai n...	pantai kuta lombok tempat tepat duduk santai...	pantai kuta lombok tepat duduk santai...	[pantai, kuta, lombok, tempat, tepat, duduk, santai...]
Pantainya sejuk, udaranya ademmmm Rekomend b...	positive	pantainya sejuk, udaranya ademmmm Rekomend b...	pantainya sejuk udaranya ademmmm rekomen b...	pantainya sejuk udaranya adem rekomendasi bang...	pantainya sejuk udara adem rekomendasi bang...	pantai sejuk udara adem rekomendasi banget baw...	[pantai, sejuk, udara, adem, rekomendasi, banget...]

3.4. Data Transformation

The data obtained from the pre-processing results was then transformed using the TF-IDF method. At this stage, weighting is carried out on the words that appear in each document. The data transformation results can be seen in the following image (Figure 2).
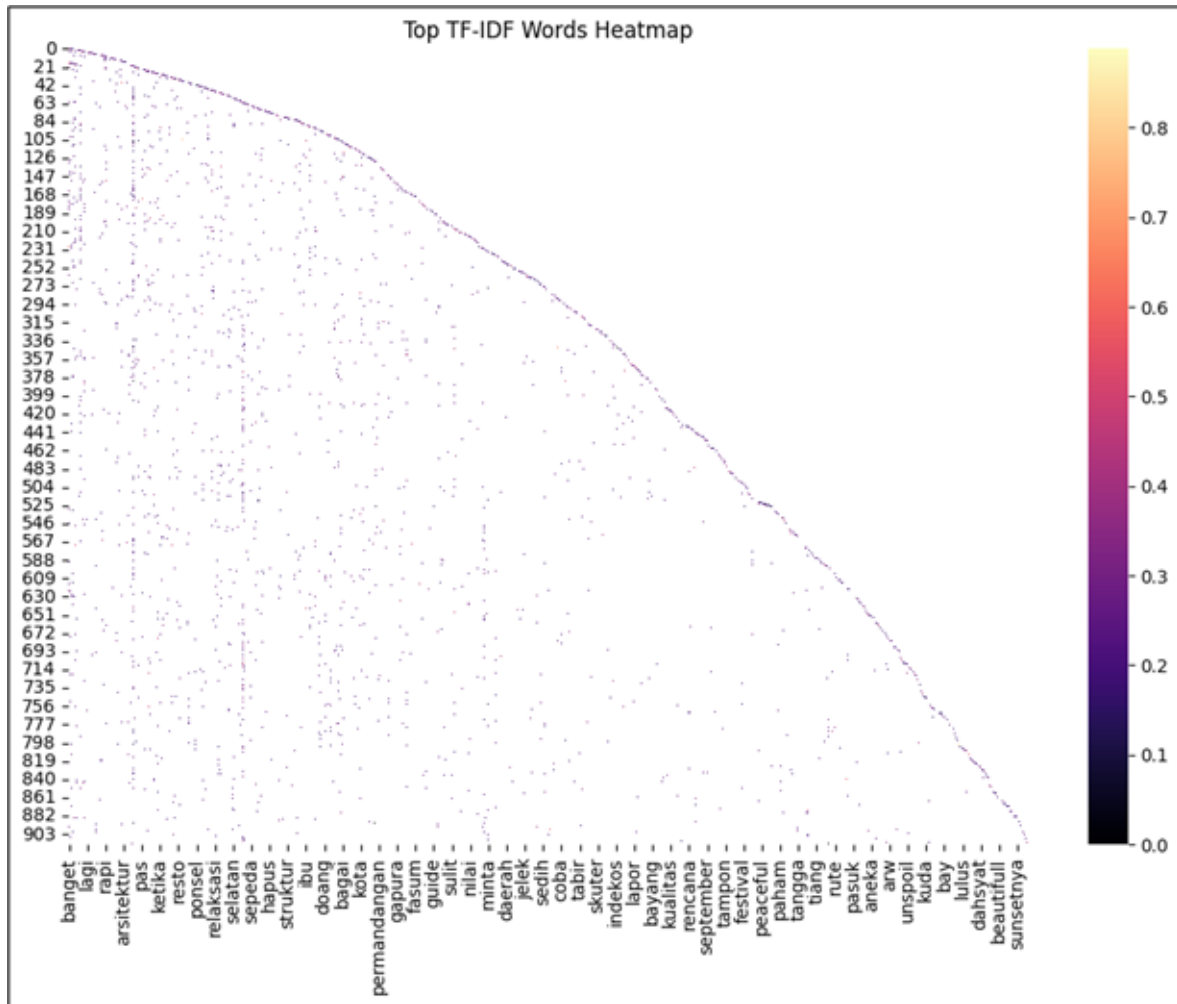

Figure 2. Data Transformation Results

The image above displays the distribution of the TF-IDF values of the top words across the document in grid form, where the color intensity indicates the magnitude of the TF-IDF values. Lighter colors indicate higher TF-IDF values, indicating that the word is highly significant in a particular document. This visualization allows for easy identification of patterns and trends in word usage, providing deeper insight into the structure

and content of the text.

3.5. Data Classification

After carrying out data transformation, the next step is data classification. In this case, the method used is the Nave Bayes method. This classification aims to find out the class of a review, whether it is an upbeat class or a negative class.

The first step in this stage is to divide the data set into two parts, namely 80% training data and 20% testing data, using the K-Fold Cross Validation method with a value of $K = 5$. In more detail, the results of each fold can be seen in the following picture (Figure 3):
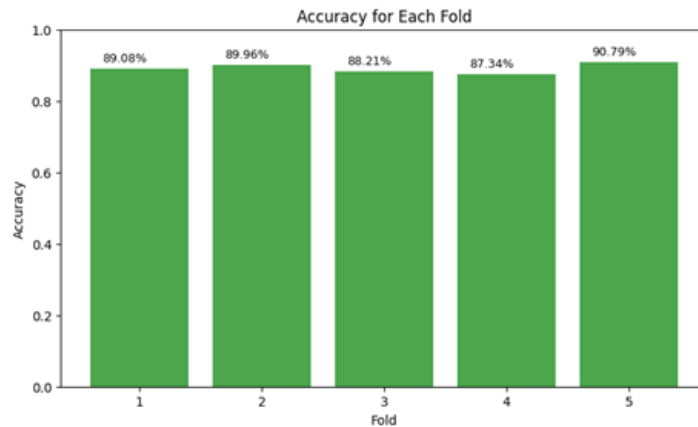

Figure 3. Accuracy Level of Each Fold

The image above shows that the level of accuracy obtained from each fold is relatively consistent, namely Fold 1: 89.08%, Fold 2: 89.96%, Fold 3: 88.21%, Fold 4: 87.34%, and Fold 5: 90.79%. This shows that the model runs stably with an average accuracy for each fold of 89.07%. Although there is a slight variation between these folds, the difference in accuracy is not very large (approximately 3.45%). This shows that the model has good generalization ability and is unaffected by slight training and testing data differences.

3.6. Data Analysis & Visualization

Based on the results of each process above, all adjectives contained in each document are taken. These adjectives represent tourists' positive and negative assessments of tourist attractions in the Mandalika Special Economic Zone. The visualization of each adjective is summarized in the following word cloud (Figure 4).

Figure 4. Word Cloud Display for Positive Review (Left) and Negative Review (Right)

The word cloud image above shows sentiment analysis of tourist attractions in the Mandalika, divided into two classes: positive and negative data. In the positive class word cloud, the most prominent words include "nice," "clean," "beautiful," "big," "quiet," and "beautiful." This shows that many visitors are satisfied with the conditions and aesthetics of tourist attractions in the Mandalika, describing them as clean, beautiful, and calming. Words such as "blue," "green," and "clear" indicate that nature and scenery in Mandalika are highly valued, especially regarding water and vegetation.

In contrast, in the negative class word cloud, the most prominent words include "the," "many," "white," "nice," and "big." Although some words such as "nice" and "big" also appear in cheerful word clouds, they can also be used in sentences expressing disappointment or dissatisfaction in a negative context. Furthermore, words such as "a little," "busy," "forced," "sad," "boxy," and "dirty" indicate unpleasant experiences, such as discomfort, lack of cleanliness, and inadequate conditions. These words suggest that although there are positive aspects, some visitors experienced things that detracted from their overall experience at the Mandalika. Overall, this sentiment analysis provides insight into the fact that although the majority of tourists have positive sentiments towards the attractions in the Mandalika SEZ, some feel dissatisfied with various aspects, such as crowds, cleanliness, or unfulfilled expectations.

4. CONCLUSION

From the analysis above, it can be concluded that a total of 1,144 reviews of tourist attractions in the Mandalika Special Economic Zone, with a division of 1,033 positive reviews and 111 negative reviews, can be classified well by the Nave Bayes algorithm model. This can be seen from the level of accuracy obtained from each fold, which is entirely consistent, namely Fold 1: 89.08%, Fold 2: 89.96%, Fold 3: 88.21%, Fold 4: 87.34%, and Fold 5: 90.79%. Although slight variations between folds exist, this confirms that the model has excellent and stable generalization abilities. Minimal variation between folds also indicates that the model is not too affected by small changes in training and testing data, which suggests a model that is not overfitting.

REFERENCES

[1] N. P. A. Aryawati, "Digital Tourism: Strategi Menarik Wisatawan Milenial ke NTB," *Jurnal Ilmiah Hospitality*, vol. 9, no. 1, pp. 87–92, Jun. 2020. doi: 10.47492/jih.v9i1.34

[2] H. Purba and I. Irwansyah, "User Generated Content dan Pemanfaatan Media Sosial Dalam Perkembangan Industri Pariwisata: Literature Review," *Professional: Jurnal Komunikasi dan Administrasi Publik*, vol. 9, no. 2, pp. 229–238, Dec. 2022. doi: 10.37676/professional.v9i2.3065

[3] N. Donthu, S. Kumar, N. Pandey, N. Pandey, and A. Mishra, "Mapping the electronic word-of-mouth (eWOM) research: A systematic review and bibliometric analysis," *Journal of Business Research*, vol. 135, pp. 758–773, Oct. 2021. doi: 10.1016/j.jbusres.2021.07.015

[4] R. Filieri, F. Acikgoz, V. Ndou, and Y. Dwivedi, "Is TripAdvisor still relevant? The influence of review credibility, review usefulness, and ease of use on consumers continuance intention," *International Journal of Contemporary Hospitality Management*, vol. 33, no. 1, pp. 199–223, Jan. 2021. doi: 10.1108/IJCHM-05-2020-0402

[5] I. W. B. Suryawan, N. W. Utami, and K. Q. Fredlina, "Analisis Sentimen Review Wisatawan pada Objek Wisata Ubud Menggunakan Algoritma Support Vector Machine," *Jurnal Informatika Teknologi dan Sains (Jinteks)*, vol. 5, no. 1, pp. 133–140, Feb. 2023. doi: 10.51401/jinteks.v5i1.2242

[6] Q. Jiang, C.-S. Chan, S. Eichelberger, H. Ma, and B. Pikkemaat, "Sentiment analysis of online destination image of Hong Kong held by mainland Chinese tourists," *Current Issues in Tourism*, vol. 24, no. 17, pp. 2501–2522, Sep. 2021. doi: 10.1080/13683500.2021.1874312

[7] Y. I. Mahendra, "Dampak Pembangunan Kawasan Ekonomi Khusus Mandalika (KEK) Terhadap Pengembangan Usaha Mikro Kecil Dan Menengah Di Kuta Lombok Tengah Di Tinjau Dalam Perspektif Ekonomi Islam," *Jurnal Econetica: Jurnal Ilmu Sosial, Ekonomi, dan Bisnis*, vol. 2, no. 2, pp. 13–21, Nov. 2020. doi: 10.0602/econetica.v2i2.104

[8] S. Damayanti, "Dampak Pembangunan Sirkuit Kuta Mandalika Terhadap Pengembangan Usaha Mikro Kecil Dan Menengah Di Tinjau Dalam Perspektif Ekonomi Islam," *Journal of Sharia Economy and Islamic Tourism*, vol. 2, no. 1, pp. 20–27, Apr. 2022. doi: 10.31764/jseit.v2i1.8345

[9] Y. A. Wahyudin, A. M. Munir, and K. Rizki, "Pemberdayaan Masyarakat di Kawasan Ekonomi Khusus (KEK) Mandalika Provinsi Nusa Tenggara Barat Melalui Indikator Pembangunan Manusia," *Prosiding Semnaskom-Unram*, vol. 4, no. 1, pp. 226–234, Dec. 2022.

[10] B. R. T. Yunarni and A. Haris, "Pemberdayaan Perekonomian Masyarakat Melalui Pemberdayaan Usaha Mikro Kecil Dan Menengah (UMKM) Di Kawasan Ekonomi Khusus (KEK) Mandalika Lombok." *JISIP (Jurnal Ilmu Sosial dan Pendidikan)*, vol. 4, no. 3, pp. 333–343, Jul. 2020. doi: 10.58258/jisip.v4i3.1224

[11] E. Sopian, R. H. Sayuti, and A. Evendi, "Model Pemberdayaan Masyarakat di Kawasan Ekonomi Khusus Mandalika Kabupaten Lombok Tengah," *Prosiding Seminar Nasional Mahasiswa Sosiologi*, vol. 1, no. 1, pp. 171–180, Apr. 2023.

[12] M. A. Satrio, "Upaya Pemerintah Indonesia dalam Meningkatkan Pariwisata Mandalika Melalui Kerangka Branding Wonderful Indonesia," *Indonesian Perspective*, vol. 6, no. 1, pp. 65–85, Mar. 2021.

[13] Z. A. Ramdani, "Peran Pemerintah Dalam Pengembangan Kawasan Ekonomi Khusus Mandalika Provinsi Nusa Tenggara Barat," *Jurnal Planoearth*, vol. 5, no. 1, pp. 1–6, May 2020. doi: 10.31764/jpe.v5i1.1639

[14] D. S. Utami and A. Erfina, "Analisis Sentimen Objek Wisata Bali Di Google Maps Menggunakan Algoritma Naive Bayes," *J-SAKTI (Jurnal Sains Komputer dan Informatika)*, vol. 6, no. 1, pp. 418–427, Mar. 2022. doi: 10.30645/j-sakti.v6i1.456

[15] Y. A. Singgalen, "Analisis Sentimen Wisatawan Melalui Data Ulasan Candi Borobudur di Tripadvisor Menggunakan Algoritma Nave Bayes Classifier," *Building of Informatics, Technology and Science (BITS)*, vol. 4, no. 3, p. 13431352, Dec. 2022. doi: 10.47065/bits.v4i3.2486

[16] N. L. W. S. R. Ginantra, C. P. Yanti, G. D. Prasetya, I. B. G. Sarasvananda, and I. K. A. G. Wiguna, "Analisis Sentimen Ulasan Villa di Ubud Menggunakan Metode Naive Bayes, Decision Tree, dan K-NN," *Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI)*, vol. 11, no. 3, pp. 205–215, Dec. 2022. doi: 10.23887/janapati.v11i3.49450

[17] A. R. Ismail and R. B. F. Hakim, "Implementasi Lexicon Based untuk Analisis Sentimen Dalam Menentukan Rekomendasi Pantai di DI Yogyakarta Berdasarkan Data Twitter: Implementasi Lexicon Based," *Emerging Statistics and Data Science Journal*, vol. 1, no. 1, pp. 37–46, Jan. 2023. doi: 10.20885/esds.vol1.iss.1.art5

[18] M. U. Albab, Y. K. P, and M. N. Fawaiq, "Optimization of the Stemming Technique on Text Preprocessing President 3 Periods Topic," *Jurnal Transformatika*, vol. 20, no. 2, pp. 1–12, Jan. 2023. doi: 10.26623/transformatika.v20i2.5374

[19] N. L. P. Merawati, A. Z. Amrullah, and I. Ismarmiaty, "Analisis Sentimen dan Pemodelan Topik Pariwisata Lombok Menggunakan Algoritma Naive Bayes dan Latent Dirichlet Allocation," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 5, no. 1, pp. 123–131, Feb. 2021. doi: 10.29207/resti.v5i1.2587

[20] P. M. R. C. Dinatha and N. A. Rakhmawati, "Komparasi Term Weighting dan Word Embedding pada Klasifikasi Tweet Pemerintah Daerah," *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, vol. 9, no. 2, pp. 155–161, May 2020. doi: 10.22146/jnteti.v9i2.90